

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:50:51 ; Search time 170 Seconds

(without alignments)  
45.183 Million cell updates/sec

Title: US-10-018-964-3\_copy\_1\_15

Perfect score: 87

Sequence: 1 YGSPSGARRRNCYE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	319	1 MYOD_HUMAN	P15172 homo sapien
2	87	100.0	320	2 O75321	O75321 homo sapien
3	84	96.6	150	2 O77799	O77799 bos taurus
4	84	96.6	318	2 O7Y882	O7Y882 bos taurus
5	84	96.6	319	1 MYOD_PIG	P49811 sus scrofa
6	84	96.6	319	1 MYOD_SHEEP	P29331 ovis aries
7	84	96.6	318	1 MYOD_MOUSE	P10085 mus musculus
8	63	72.4	318	1 MYOD_RAT	O02346 rattus norv
9	63	72.4	318	2 O8C6B1	O8C6B1 mus musculus
10	60	69.0	297	2 MYOD_COTJA	P15752 coturnix co
11	60	69.0	297	2 O6DV59	O6DV59 melospiza g
12	60	69.0	298	1 MYOD_CHICK	P13904 gallus gall
13	57	65.5	289	1 MYOD_XENTLA	O8AVZ0 xenopus lae
14	57	65.5	289	2 O8AVZ0	O8AVZ0 xenopus lae
15	50	57.5	172	2 O75V41	O75V41 physcomitre
16	50	57.5	417	2 O75V42	O75V42 physcomitre
17	49.5	56.9	288	2 O7T109	O7T109 xenopus tro
18	48	55.2	652	2 O8W8M9	O8W8M9 dugesia tlg
19	48	55.2	652	2 O8W8M9	O8W8M9 dugesia tlg
20	47	54.0	271	1 FRA1_HUMAN	P15407 homo sapien
21	47	54.0	273	1 FRA1_MOUSE	P48755 mus musculus
22	47	54.0	273	2 O7TMD4	O7TMD4 mus musculus
23	47	54.0	275	1 FRA1_RAT	P10158 rattus norv
24	47	54.0	652	2 O8W8X1	O8W8X1 dugesia tlg
25	47	54.0	2340	2 O7RSD0	O7RSD0 plasmodium
26	46	52.9	932	2 O9U966	O9U966 eimeria ten
27	45.5	52.3	932	2 O7X893	O7X893 oryza sativ
28	45	51.7	120	2 O9ZMY5	O9ZMY5 bacterioph
29	45	51.7	139	2 O6ZAR3	O6ZAR3 oryza sativ
30	45	51.7	1184	1 DP3A_MYCBO	P61978 mycobacteri
31	45	51.7	1184	1 DP3A_MYCTU	P61977 mycobacteri

#### ALIGNMENTS

32	44	50.6	128	2	O05400	O05400 streptomyc
33	44	50.6	257	2	O7Q307	O7Q307 anopheles g
34	44	50.6	278	2	O7ZUG9	O7ZUG9 brachydanio
35	44	50.6	287	1	MF25_XENTLA	P16076 xenopus lae
36	44	50.6	288	2	O6GN48	O6GN48 xenopus lae
37	44	50.6	568	2	O84W52	O84W52 xenopus lae
38	44	50.6	568	2	O9LFB9	O9LFB9 arabidopsis
39	44	50.6	949	2	O9A9J5	O9A9J5 caulobacter
40	44	50.6	1038	2	O8PAY6	O8PAY6 xanthomonas
41	43	49.4	99	2	O6LHG2	O6LHG2 photobacter
42	43	49.4	134	2	O8W8M2	O8W8M2 caenorhabdi
43	43	49.4	153	2	O8GJ48	O8GJ48 chlorobium
44	43	49.4	153	2	O8GJ49	O8GJ49 chlorobium
45	43	49.4	154	2	O8GJ35	O8GJ35 chlorobium

RESULT 1  
ID MYOD\_HUMAN STANDARD; PRT; 319 AA.

AC P15172;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Myoblast determination protein 1 (Myogenic factor 3) (Myf-3).  
GN Name=MYOD1; Synonyms=MYF3, MYOD,  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91212198; PubMed=1850513;  
RA Pearson-White S.H.;  
RT "Human Myod: cDNA and deduced amino acid sequence."  
RL Nucleic Acids Res. 19:1148-1148 (1991).  
[2]  
RP SEQUENCE OF 53-319 FROM N.A.  
RC TISSUE=skelatal muscle;  
RX MEDLINE=90059960; PubMed=2583111;  
RA Braun T., Bober S., Buschhausen-Denker G., Kotz S., Grzeschik K.-H.,  
RA Arnold H.H.;  
RT "Differential expression of myogenic determination genes in muscle  
cells: possible autoactivation by the Myf gene products."  
RL EMBO J. 8:3617-3625 (1989).  
[3]  
RP REVIEW ABOUT ACETYLATION AND DEACETYLATION.  
RX MEDLINE=21423399; PubMed=11532390; DOI=10.1016/S0959-437X(00)00224-0;  
RA McKinsey T.A., Zhang C.L., Olson E.N.;  
RT "Control of muscle development by duoling HMTs and HDACs."  
RL Curr. Opin. Genet. Dev. 11:497-504 (2001).  
CC -I- FUNCTION: Involved in muscle differentiation (myogenic factor).  
CC Induces fibroblasts to differentiate into myoblasts. Activates  
CC muscle-specific promoters. Interacts with and is inhibited by the  
CC twist protein. This interaction probably involves the basic  
CC domain of both proteins (By similarity).  
CC -I- SUBUNIT: Efficient DNA binding requires dimerization with another  
CC bHLH protein. Seems to form active heterodimers with ITP-2.  
CC -I- SUBCELLULAR LOCATION: Nuclear.  
CC -I- PTM: Acetylated by a complex containing Epi300 and PCAF. The  
CC acetylation is essential to activate target genes. Conversely, its  
CC deacetylation by SIRT1 inhibits its function (By similarity).  
CC -I- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

CC -----  
 DR EMBL, X56677, CAA40000.1; -  
 DR EMBL, X17650, CAA35640.1; -  
 DR PIR, S26827, S26827.  
 DR HSSP, P10085, IMDY.  
 DR TRNSPAC, T00519; -  
 DR TRNSPAC, T00525; -  
 DR Genem, HENC:7611; MYOD1.  
 DR MIM, 159970; -  
 DR GO, GO:0005634; C:nucleus; TAS.  
 DR GO, GO:0003705; F:RNA polymerase II transcription factor act...; TAS.  
 DR GO, GO:0003713; F:transcription coactivator activity; TAS.  
 DR GO, GO:0007519; P:myogenesis; TAS.  
 DR GO, GO:0006468; P:protein amino acid phosphorylation; TAS.  
 DR GO, GO:0006357; P:regulation of transcription from Pol II promoter; TAS.  
 DR InterPro, IPR002546; Basic.  
 DR InterPro, IPR001092; HLH\_Basic.  
 DR Pfam, PF00010; HLH, 1.  
 DR Pfam, PF01586; Basic, 1.  
 DR PROSITE, PS50868; HLH, 1.  
 DR Acetylation; Developmental protein; Differentiation; DNA-binding;  
 KW Myogenesis; Nuclear protein; transcription regulation.  
 FT DNA\_BIND 109 121 Basic motif.  
 FT DOMAIN 122 161 Helix-loop-helix motif.  
 FT CONFLICT 124 124 K -> E (in Ref. 2).  
 SO SEQUENCE 319 AA, 34490 MW, AAE93C1A54318770 CRC64;

QY 1 YSGPSPGARRRNCYE 15  
 DB 212 YSGPSPGARRRNCYE 226

Query Match 100.0%; Score 87; DB 1; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;